Amendments to the Specification:

Please insert the following paragraph following the title of the invention on page 1 of the specification:

-- This is a continuation of U.S. Patent Application Serial No. 09/230,158, filed January 19, 1999, which is the National Stage of International Application PCT/IB97/00981, filed July 25, 1997, which claims priority to DE19630390.7, filed July 26, 1996. Each of these documents is incorporated by reference herein in their entireties.--

Please amend the paragraph at page 1, lines 2-14 of the specification as follows:

-- The present invention relates to novel proteins, in particular membrane proteins or proteins which are firmly associated with the membrane, which are derived from *Helicobacter pylori* (*H. pylori*) and which contain one of the peptide sequences selected from SEQ ID NO:1, 2, 3, 6, 10, 11, 12, 14, 15, 16, 17, 18, or 19 according to Tables 1a1c, or to parts or homologues thereof having a minimum length of five amino acids, and to their preparation and use as pharmaceutical compositions, in particular as vaccines, or as a diagnostic agent. Based on these data, genes coding for these and related proteins were also isolated as shown in SEQ ID NOs: 20, 21, 22, 23, 24, 25, 26,~~and~~ 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, and 37.--

Please amend the paragraph at page 7, line 33 to page 8, line 4 of the specification as follows:

-- In a twelfth aspect of the present invention, there is provided a protein from *H. pylori* containing one of the peptide sequences deduced from SEQ ID NO:21, 22, 23, 24, 25,

26, ~~and~~ 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, and 37,or parts or homologues thereof having a minimum length of five amino acids.--

Please amend the paragraph at page 8, lines 5-9, of the specification as follows:

-- In a thirteenth aspect of the present invention there is provided a peptide having the amino acid sequence deduced from SEQ ID NO:21, 22, 23, 24, 25, 26, ~~or~~ 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, or 37, or parts or homologues thereof having a minimum length of five amino acids.--

Please amend the paragraph at page 8, lines 10-15 of the specification as follows:

-- In a fourteenth aspect of the present invention there is provided a peptide selected from the C-terminal region of the peptide sequence of SEQ ID NO:20 or homologue thereof. Preferably said peptide is selected from RDPKFNLAHIEKEFEVWNWDYRA (SEQ ID NO:51) and EKHQKMMKDMHGKDMHHTKKKK (SEQ ID NO:52), or parts or homologues thereof.--

Please amend the paragraph at page 16, line 31 to page 17, line 5 of the specification as follows:

-- Using oligonucleotides deduced from the N-terminal sequences of SEQ ID NOs: 5, 7, 8, 10, 12 and 15, the genes corresponding to the SEQ ID NOs:5, 8, 10, 12 and 15 were isolated and are specified as SEQ ID NOs: 20 and 21 (catalase),~~24~~ 30 and 31 (50 kD membrane protein), ~~25~~ 32 and 33 (42 kD protein), ~~26~~ 34 and 35 (36/35/32 kD protein) and ~~23~~ 28 and 29 (28 kD protein). The gene coding for HopC could not be isolated using oligonucleotide 7. However, oligonucleotide 7 hybridizes with an homologous gene

specified as ~~SEQ ID NO:21~~ SEQ ID NOs:24 and 25 (HopX). Two additional genes which belong to this family were isolated and are specified as ~~SEQ ID NO:21~~ SEQ ID NOs:22 and 23 (HopY) and ~~SEQ ID NO:22~~ SEQ ID NOs:26 and 27 (HopZ).--

Please amend the paragraph at page 17, lines 6-9 of the specification as follows:

-- Another approach is given by the recent access to the complete genomic sequence of *H. pylori* on the internet which allowed, for example, the identification of ~~SEQ ID NO:27~~ SEQ ID NOs:36 and 37.--

Please amend the paragraph at page 20, line 20 to page 21, line 9 of the specification as follows:

-- For example, according to the present invention, a DNA vaccine can be prepared on the basis of the polynucleotides, or a diagnostic agent can be prepared on the basis of the polymerase chain reaction (PCR diagnosis), or an immunotest, for example a Western blot test or an enzyme immunotest (ELISA) can be prepared on the basis of the antibodies. Furthermore, the novel proteins or peptides, or their immunogenic moieties, in particular when they contain an uninterrupted sequence of unambiguously determined amino acids, having a minimum length of five amino acids, preferably six amino acids and, in particular, in the case of the novel peptides having the SEQ ID NOs:1, 2, 3, 6, 10, 11, 12, 14, 15, 16, and 19 and peptides or proteins encoded by the DNA sequences of SEQ ID NOs:20, 21, 22, 23, 24, 25, 26, ~~and~~ 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, and 37, at least ten consecutive amino acids, can be used as antigens for immunizing mammals. In this context, the two C-terminal regions C1 and C2 specific for *H. pylori* catalase (c.f. Example 6) can also be used as immunogens. The antibodies which are formed by the immunization, or antibodies which are

prepared by means of recombinant DNA methods (see, for example, Winter G. & Milstein C. (1991) Nature, 293-299, Vol. 349) can, inter alia, prevent adhesion of the bacteria to the mucosal surface, attract macrophages for the purpose of eliminating bacteria, and activate the complement system for the purpose of lysing the bacteria.--

Please amend the paragraph at page 25, lines 5-40 of the specification as follows:

-- The resulting oligonucleotides were deduced from the resulting N-terminal sequences of SEQ ID NOs:5, 7, 8, 10, 12 and 15:

SEQ ID NO:	Oligonu-cleotide	Amino acid sequence and predicted nucleotide
5	1	Val Asn Lys Asp Val Lys Gln Thr Xaa GTI AAT AAA GAT GTI AAA CAA ACT TGT C C Ala Phe Gly Ala Pro (SEQ ID NO:45) GCI TTT GGC GCI CCT (SEQ ID NO:38)
7	2	Gly Gly Phe Phe Thr Val Gly Tyr Gln Leu GGC GGC TTT TTT ACT GTG GGC TAT CAA TTA C G Gly Gln Val Met Gln (SEQ ID NO:46) GGC CAA GTG ATG CAA (SEQ ID NO:39)
8	3	(Val) (Thr) Tyr Glu Val His (Gly) Asp Phe Ile GTG ACT TAT GAA GTG CAT GGC GAT TTT ATC C T Asn Phe (Ser) Lys Val (SEQ ID NO:47) AAT TTT AGC AAA GT (SEQ ID NO:40) C
10	4	Lys Glu Lys Phe Asn Arg Thr Lys Pro (SEQ ID NO:48) AAA GAA AAA TTT AAC AGA ACC AAA CCT (SEQ ID NO:41) T
12	5	Glu Lys Asn Gly Ala Phe Val Gly Ile Ser GAA AAA AAT GGI GCI TTT GTG GGC ATT AGC C Leu Glu Val Gly Arg Ala Asp Gln Lys (SEQ ID NO:49) TTI GAG GTT GGI AGA GCT GAT CAA AAA (SEQ ID NO:42)
15	6	Trp Ser Ala Ala Phe Val Gly Val Asn TGG AGC GCT GCT TTT GTG GGC GTG AAT Tyr Gln Val Ser Met Ile Gln Asn Gln Thr TAT CAA GTG AGC ATG ATT CAA AAT CAA ACT C C Lys Met Val Asn Asp (SEQ ID NO:50)

		AAA ATG GTG AAT GAT (SEQ ID NO:43)

Please amend the paragraph at page 27, line 16 to page 28, line 5 of the specification as follows:

-- SEQ ID NO:20 describes the DNA sequence which encodes the catalase of*H. pylori*. The nucleotide region 337 to 378 describes the hybridization site with oligonucleotide 1. The catalase gene of *H. pylori* has been described in 1996 by Stefan Odenbreit, Björn Wieland and Rainer Haas (J. Bacteriol. 178, 6960-6967) and is therefore not new. However, when comparing the amino acid sequences of the catalases of *Escherichia coli*,*Bacillus firmus*, *B. subtilis* A, *B. subtilis* B, rats, mice, cattle, humans, *Staphylococcus violaceus*, *Haemophilus influenzae*, *B. fragilis*, *Pseudomonas mirabilis*, *B. pertussis*, and *P. syringae* with the amino acid sequence of *H. pylori*, it is possible to identify two C-terminal regions of C1 (RDPKFNLAHIEKEFEVWNWDYRA; SEQ ID NO:51) and C2 (EKHQKMMKDMHGKDMHHTKKKK; SEQ ID NO:52), which are specific to *H. pylori* catalase. These two peptides were synthesized using standard techniques, coupled to KLH and used for immunizing rabbits. These rabbits developed antibodies against the two peptides, which reacted in the Western Blot analysis with *H. pylori* catalase which had been produced by recombinant technique. These *H. pylori* catalase-specific regions may conceivably be used for developing a vaccine which avoids the problem complex of autoimmune reactions or for the development of a diagnostic which reacts specifically with *H. pylori* catalase.--

Please amend the paragraph at page 28, lines 6-13 of the specification as follows:

-- ~~SEQ ID NO: 21~~ SEQ ID NO:22 describes a nucleotide sequence which was identified by hybridization with the oligonucleotide 2. The oligonucleotide hybridized with the sequence of nucleotide 1240 to 1284. This encodes a sequence which is homologous to the porin Hop C (Exner et al., 1995) and is identical with the published amino-terminal sequence EDDGGFFTVGYQLGQVMQDVQNPG (SEQ ID NO:44) in positions 1, 2, 3, 4, 9, 10, 11, 12, 14, 18 and 22.--

Please amend the paragraph at page 28, lines 14-22 of the specification as follows:

-- The porins Hop A, Hop B, Hop C, and Hop D have identical amino acids in 9 positions of the 20 N-terminal amino acids (Exner et al., 1995). In 8 of these positions, there are identical positions also in the sequence described ~~in the present publication~~ herein; in the 9th position, a conserved amino acid exchange is present (Val~~→~~ Ile). It can thus be assumed that the protein described ~~in the present publication~~ herein is equally part of this group of the porins; it was therefore termed HopX.--

Please amend the paragraph at page 29, lines 5-12 of the specification as follows:

-- ~~SEQ ID NO:22~~ SEQ ID NO:26 describes a nucleotide sequence which was concomitantly isolated and sequenced during the screening process. The amino acid sequence deduced encodes the 392 C-terminal residues of a protein which shows a high homology with HopX (33% identity) and HopY (28%) identity and which was therefore termed HopZ. The gene region which encodes the N-terminal portion of the protein is currently being isolated.--

Please amend the paragraph at page 29, lines 13-32 of the specification as follows:

-- ~~SEQ ID NO:23~~ SEQ ID NO:28 describes a DNA sequence which encodes a hitherto undescribed protein. The nucleotide region 696 to 767 describes the hybridization site with the oligonucleotide 6. On the basis of the N-terminal protein sequence which has been determined, in which it was not possible unequivocally to determine the amino acids in the first two positions, and on the basis of the hydrophobicity of the N-terminal protein sequence deduced from the nucleic acid sequence, it can be concluded that the protein deduced has a signal sequence of 17 amino acids. The mature protein of 231 amino acids has a molecular weight of 26.4 kD and an isoelectric point of 10.3. Thus, the molecular weight is quite close to the molecular weight of 28 kD which has been determined by SDS gel electrophoresis. The amino acid sequence deduced is homologous with the sequences of the proteins HopX, HopY and HopZ, for which the GCG Bestfit Programme determined the identity values of 41%, 38% and 41%, respectively. The 28 kD protein thus also seems to be part of the family of the porins or porin-like proteins.--

Please amend the paragraph at page 29, line 33 to page 30, line 21 of the specification as follows:

-- ~~SEQ ID NO:24~~ SEQ ID NO:30 describes a DNA sequence which encodes the non-heat-modifiable 50 kD membrane protein. This protein was first characterized by Exner et al., 1995, and an N-terminal sequence of the protein was determined. Using the approach described by us, we were then able to describe, with SEQ ID NO:8, an N-terminal sequence which is identical to the sequence described by Exner et al. (1995). With the aid of the oligonucleotide 3, which had been deduced using the method illustrated in Example 5 and had been used for screening a *H. pylori* gene library using the above-described methods, it was then possible to identify a DNA fragment which encodes the 50 kD membrane protein.

Using other standard methods, it was then possible to determine the nucleic acid sequence described in ~~SEQ ID NO:24~~ SEQ ID NO:30, which encodes a mature protein of 499 amino acids which has a molecular weight of 56.3 kD and an isoelectric point of 9.75. Due to the data of the N-terminal sequencing procedures and the hydrophobicity of the N-terminal sequence, a signal sequence of 29 amino acids is assumed. The amino acid residues 236 to 254 contain a hydrophobic region which is large enough to act as a transmembrane region. Based on such data and using standard methods for epitope analysis, it is possible to identify regions which might be presented on the surface of the bacteria. Such regions might be used for developing a vaccine or a diagnostic.--

Please amend the paragraph at page 30, line 22 to page 31, line 1 of the specification as follows:

-- ~~SEQ ID NO:25~~ SEQ ID NO:32 describes a DNA sequence 2825 bp in size which was identified by means of hybridization with oligonucleotide 4, which was deduced from SEQ ID NO:10. Oligonucleotide 4 hybridized with the nucleotide region 897 to 923 of the described sequence of ~~SEQ ID NO:25~~ SEQ ID NO:32. The protein has no signal sequence. The encoding region of ~~SEQ ID NO:25~~ SEQ ID NO:32 codes for a protein of 399 amino acids with a molecular weight of 43.6 kD and an isoelectric point of 5.0. A search for homologous sequences using the BLASTP program (S.F. Altschul et al., 1990, J. Mol. Biol. 215, 403-410) identified the 42 kD antigen of *H. pylori* as the elongation factor TU. The maximum percentage of identity (89%) was found with the elongation factor TU from *Wolinella succinogens* (W. Ludwig et al., 1993 Antonie van Leeuwenhoek 64, 285-305).

Please amend the paragraph on page 31, lines 2-18 of the specification as follows:

-- ~~SEQ ID NO:26~~ SEQ ID NO:34 describes a DNA sequence 2182 bp in size which hybridizes with oligonucleotide 5, which had been deduced from SEQ ID NO:12. Oligonucleotide 5 hybridized with a *Sau*3AI fragment (position 1 to 575) of the gene library starting from position 524. The screening of different DNA libraries with specific oligonucleotides allowed the isolation of the complete gene described in ~~SEQ ID NO:26~~ SEQ ID NO:34. An amino acid sequence which is identical to the one from SEQ ID NO:12 can be deduced from ~~SEQ ID NO:26~~ SEQ ID NO:34. Both protein sequencing and the hydrophobicity of the N-terminal sequence deduced allow the conclusion that the antigen has a signal sequence. The mature protein consists of 328 amino acid residues with a molecular weight of 36.1 kD and an isoelectric point of 9.95. No homologous proteins were identified using the BLASTP program (S.F. Altschul et al., 1990).--

Please amend the paragraph at page 31, lines 19-26 of the specification as follows:

-- The sequences described in ~~SEQ ID NOS:20 to 26~~ SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, and 34 indicate nucleotide sequences which encode antigens of the *H. pylori* strain ATCC 43504. However, it is known for *H. pylori* that heterogeneity between identical antigens may exist among various strains. We therefore claim not only the sequences described in ~~SEQ ID NOs:21-26~~ SEQ ID NOs:23, 25, 27, 29, 31, 33, and 35, but in addition also the sequences of other *H. pylori* strains which are homologous with the sequences described herein.--

Please amend the paragraph on page 31, line 31 to page 32, line 34 as follows:

-- The Institute for Genomic Research (TIGR) released the DNA sequence from *H. pylori* on 24th June 1997. This new information ~~can be accessed on the internet at~~

~~www.tigr.org~~ is available from TIGR. Using the TBLASTN program (Altschul et al., 1997, Nucleic Acids Research 25, ~~in press~~, 3389-3402) the peptide sequences listed in Tables 1a – 1c can be aligned to amino acid sequence data deduced from all six reading frames of the *H. pylori* strain 26695. Having access to the genomic DNA sequence, DNA sequences corresponding to the aligned amino acid sequences can be identified using GCG (Genetic Computer Group) programs. This approach is shown for SEQ ID NO:19, for example. The sequence of SEQ ID NO:19 is aligned with a very similar sequence using the TBLASTN program. ~~SEQ ID NO:27 describes~~ SEQ ID NOs: 36 and 37 describe the nucleic acid sequence and deduced amino acid sequence from the coding region of a *H. pylori* gene (strain 26695) localised between position 843212 and 843691 of the genomic sequence. The protein has no signal sequence. The N-terminal sequence of SEQ ID NO:19 is highly homologous to the N-terminal region of the deduced amino acid sequence from amino acid residue 1 to 15. Only one different amino acid residue is present at position 4: the nucleotide sequence found by the alignment encodes a Ser residue in this position instead of an Asn residue determined by N-terminal sequencing. This can be explained by strain specific differences. The identified nucleic acid sequence in ~~SEQ ID NO:27~~ SEQ ID NO:36 codes for a protein of 159 amino acid residues with a molecular weight of 18.2 kD and an isoelectric point of 7.2. The molecular weight is very close to that of 17 kD determined from SDS polyacrylamide gel electrophoresis. A search for homologous sequences using the BLASTP program (S.F. Altschul et al., 1990) shows that the 17 kD antigen is very homologous to "hydroxymyristol-[acyl carrier protein] dehydratase" from different bacteria.--